From minimal signed circuits to the dynamics of Boolean regulatory networks

Elisabeth Remy, Paul Ruet

BIOINFORMATICS

Vol. 00 no. 00 2008
Pages 1–7

From minimal signed circuits to the dynamics of Boolean regulatory networks

Elisabeth Remy, Paul Ruet

CNRS, Institut de Mathématiques de Luminy
Campus de Luminy, Case 907, 13288 Marseille Cedex 9 (France)

ABSTRACT

It is acknowledged that the presence of positive or negative circuits in regulatory networks such as genetic networks is linked to the emergence of significant dynamical properties such as multistability (involved in differentiation) and periodic oscillations (involved in homeostasis). Rules proposed by the biologist R. Thomas assert that these circuits are necessary for such dynamical properties. These rules have been studied by several authors. Their obvious interest is that they relate the rather simple information contained in the structure of the network (signed circuits) to its much more complex dynamical behaviour. We prove in this article a non-trivial converse of these rules, namely that certain positive or negative circuits in a regulatory graph are actually *sufficient* for the observation of a restricted form of the corresponding dynamical property, differentiation or homeostasis. More precisely, the crucial property that we require is that the circuit be *globally minimal*. We then apply these results to the vertebrate immune system, and show that the 2 minimal functional positive circuits of the model indeed behave as *modules* which combine to explain the presence of the 3 stable states corresponding to the Th0, Th1 and Th2 cells.

Contact: ruet@iml.univ-mrs.fr

1 INTRODUCTION

The activity of a biological cell is to a large extent controlled by genetic regulation, an interactive process usually represented by graphs called genetic regulatory networks: in these graphs, vertices denote genes or regulatory products (e.g., RNA, proteins) and edges denote regulatory interactions between these genes or their products [5, 23, 28]. These regulatory interactions are further directed and signed ($+1$ or -1) to denote activatory versus inhibitory effects.

In order to relate regulatory networks to relevant dynamical properties, biologists often use them as a basis to generate dynamical models, using either a differential framework or a discrete framework [5]. The biological pertinence of the model considered is then evaluated by comparing numerical simulations with experimental observations, for instance biochemical characterizations of cellular states, phenotypes of genetic mutants, etc.

Since the computational complexity of these simulations is, in general, exponentially increasing with the size of the network, some mathematical properties could fruitfully help in controlling the space of necessary simulations. In the early 1980's, the biologist R. Thomas proposed two simple rules relating the structure of regulatory networks to their dynamical properties [30]:

1. a necessary condition for multistability (i. e., the existence of several stable fixed points in the dynamics) is the existence of a positive circuit in the regulatory network (the sign of a circuit being defined as the product of the signs of its edges);
2. a necessary condition for the existence of an attractive cycle in the dynamics is the existence of a negative circuit.

These two types of dynamical properties correspond to important biological phenomena: cell differentiation processes in the first case, homeostasis such as stable periodic behaviours (e.g., cell cycle or circadian rhythms) in the second case. Several authors have proposed demonstrations of these rules in a differential framework [15, 25, 8, 26], and more recently in a discrete framework [3, 20, 17], in which the expression levels of genes are discretised and modelled as elements of a finite set such as $\{0, 1\}$. Discrete approaches are indeed increasingly used in biology [9, 24, 2, 6, 22] because of the qualitative nature of most experimental data, together with a wide occurrence of non-linear regulatory relationships. In [20] in particular, the dynamics of a system of n genes is represented by a map $f : \{0,1\}^n \to \{0,1\}^n$, and a signed directed graph $G(f)(x)$ is associated to each state of the system $x \in \{0,1\}^n$. This graph corresponds to a local notion of regulatory graph (as in [26] for instance), and is mathematically defined by means of the discrete Jacobian matrix $J(f)(x)$ [21]. The required definitions are recalled in Section 2.

While these results provide graphic conditions which are necessary to observe some dynamical properties, they do not give sufficient conditions at all, while biologists often acknowledge certain positive or negative circuits as responsible for some dynamical behaviour [29, 31]. In the very specific case of discrete isolated circuits however, i. e., when the regulatory graph $G(f)(x)$ does not depend on the state x and consists in a circuit, [16] provide an extensive analysis of the dynamics, recalled in Section 3.

In the present paper, we show that the presence of certain positive or negative circuits in a local graph $G(f)(x)$ suffices for the observation of the corresponding dynamical property (multistability or a restricted version of homeostasis). More precisely, the crucial property that C has to meet is to be *globally minimal*, i. e., minimal as a circuit in the global graph $G(f) = \bigcup_{x\in\{0,1\}^n} G(f)(x)$ obtained by taking the union of all local graphs. In Section 4 we define a restricted form of fixed points and attractive cycles for each set I of genes, and we show that if C is a globally minimal positive (resp. negative) circuit with vertex set $\{k_1, \ldots, k_p\}$, then a suitably defined restriction of f to $\{k_1, \ldots, k_p\}$ has two fixed points (resp. an attractive cycle). These results provide:

- a non-trivial converse to Thomas' rules in the discrete framework,
- a natural approach to the question of modularity of regulatory networks, namely: given pieces of a network for which the dynamics is known, how do they combine to produce a global (more complex) behaviour? Our results on the effect of specific functional circuits in a network gives insights into this line of research.

In Section 5, we present a biological illustration of our approach: the Th-lymphocyte differentiation in the vertebrate immune system, and we apply the results of Section 4. The analysis of globally minimal circuits enables to recover the presence of the 3 stable states, which correspond to the Th0 (naive), Th1 and Th2 cells.

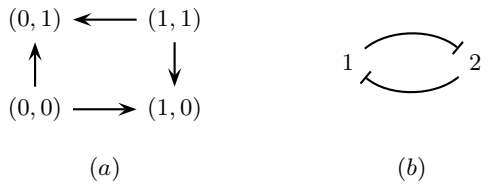

Fig. 1. (a) Asynchronous dynamics: the states of a system consisting in two variables 1 (horizontal axis) and 2 (vertical axis) are pictured; an arrow from state x to state $\overline{x}^i$ means that $f_i(x) \neq x_i$. (b) The regulatory graph $G(f)(x)$, which turns out not to depend on x. Edges represent activations or inhibitions and are respectively denoted by arrows $\rightarrow$ and T-end notation $\dashv$, which are more standard in biological literature than $\overset{+1}{\rightarrow}, \overset{-1}{\rightarrow}$.

2 BOOLEAN DYNAMICS AND DISCRETE JACOBIAN MATRICES

2.1 Notations

Let us start with preliminary notations. For $\beta \in \{0,1\}$, we define $\overline{\beta}$ by $\overline{0} = 1$ and $\overline{1} = 0$. Let n be a positive integer. For $x \in \{0,1\}^n$ and $I \subseteq \{1,\dots,n\}$, $\overline{x}^I \in \{0,1\}^n$ is defined by:

$$(\overline{x}^I)_i = \begin{cases} x_i & \text{for } i \notin I, \\ \overline{x_i} & \text{for } i \in I. \end{cases}$$

When $I = \{i\}$ is a singleton, $\overline{x}^{\{i\}}$ is denoted by $\overline{x}^i$. The distance $d : \{0,1\}^n \times \{0,1\}^n \rightarrow \{0,1,\dots,n\}$ is the Hamming distance: $d(x,y)$ is the number of $i \in \{1,\dots,n\}$ such that $x_i \neq y_i$. Suppose $0 \leqslant k \leqslant n$, and I is a k-element subset of $\{1,\dots,n\}$. Then each $x \in \{0,1\}^n$ generates an affine k-dimensional subspace $x[\![I]\!]$ of $\{0,1\}^n = \mathbb{F}_2^n$ defined by:

$$x[\![I]\!] = \{y \in \{0,1\}^n \text{ such that } y_j = x_j \text{ for all } j \notin I\}.$$

2.2 Dynamics

In the context of genetic regulatory networks, we are interested in the evolution of the system consisting of n genes, which are denoted by the integers $1,\dots,n$. We consider $\{0,1\}^n$ as the set of *states* of this dynamical system. Given a state $x = (x_1,\dots,x_n) \in \{0,1\}^n$, x_i denotes the (discretized) expression level of gene i. These expression levels are either 0 (when the gene product is considered absent or inactive) or 1 (when the gene product is present and active).

In discrete models, a dynamics is a binary relation R which we assume to be irreflexive: R gives the rule for updating a state, i. e., it is the set of pairs of states (x,y) such that state x can lead to state y. In particular, a *stable state* is a state x such that for no y, $(x,y) \in R$.

In the context considered in this paper (genetic networks), it is not realistic to assume a simultaneous update of all variables. Indeed, the Boolean dynamical systems we are interested in can be seen as discretizations of piecewise-linear differential systems [7, 30, 5, 27], and for these systems, the set of trajectories meeting more than one threshold hyperplane at a time has measure 0. We shall therefore consider *asynchronous dynamics*, i. e., relations R such that:

$$(x,y) \in R \text{ implies } d(x,y) = 1,$$

i. e., $y = \overline{x}^i$ for some i. Clearly, the asynchronous dynamics encompasses, among many others, the realistic trajectories, and a more refined analysis would take into account, e.g., delays and probabilistic issues. Such an asynchronous dynamics R may be non-deterministic (it needs not be a function), but even then, it is possible and convenient to represent it by a map $f : \{0,1\}^n \rightarrow \{0,1\}^n$ with coordinate functions $f_1,\dots,f_n$, defined by:

$$f_i(x) \neq x_i \text{ when } (x, \overline{x}^i) \in R. \tag{1}$$

Observe that a stable state is then a fixed point x for f ($f(x) = x$). More generally, if $I \subseteq \{1,\dots,n\}$, an *$I$-fixed point* is an x such that $f_i(x) = x_i$ for all $i \in I$, i. e., the coordinates in I are fixed under f.

Given such a map f, the corresponding asynchronous dynamics is defined in a straightforward way, and for each $x \in \{0,1\}^n$ and $i = 1,\dots,n$, $f_i(x)$ denotes the value to which x_i, the expression level of gene i, tends when the system is in state x.

For instance, the asynchronous dynamics corresponding to the map $f : \{0,1\}^2 \rightarrow \{0,1\}^2$ defined by $f(x) = (\overline{x_2}, \overline{x_1})$ is illustrated in Figure 1.

A *trajectory* in the dynamics is a sequence of states $(x^1,\dots,x^r)$ such that for each $i = 1,\dots,r-1$, $(x^i, x^{i+1}) \in R$, and a *cycle* is a trajectory of the form $(x^1,\dots,x^r,x^1)$ with $r \geqslant 2$. We shall be especially interested in a specific class of cycles which correspond to periodic oscillations: a cycle $(x^1,\dots,x^r,x^1)$ is said to be *attractive* when no trajectory may leave it, i. e., for all $i = 1,\dots,r$, $d(x^i, f(x^i)) = 1$. More generally, if $I \subseteq \{1,\dots,n\}$, a cycle $(x^1,\dots,x^r,x^1)$ is said to be *I-attractive* when for all $i = 1,\dots,r$, by considering indices modulo r:

- the only coordinate $\varphi(i)$ such that $x^{i+1} = \overline{x^i}^{\varphi(i)}$ belongs to I,
- the set J such that $f(x^i) = \overline{x^i}^{J \cup \{\varphi(i)\}} = \overline{x^{i+1}}^J$ is disjoint from I.

Figure 2 shows an example of dynamics with two attractive cycles:

$$((0,0),(1,0),(0,0)) \text{ and } ((0,1),(1,1),(0,1)).$$

We shall see examples of I-attractive cycles in Section 5.

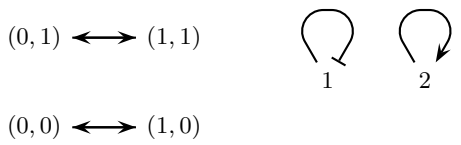

Fig. 2. A dynamics with no fixed point but a positive loop in the (constant) regulatory graph. The notation is the same as in Figure 1.

2.3 Discrete Jacobian matrices and signed directed graphs

Given $f : \{0,1\}^n \to \{0,1\}^n$, we attach to each $x \in \{0,1\}^n$ its *discrete Jacobian matrix* $J(f)(x)$ as defined in [21]: $J(f)(x)$ is the $n \times n$ matrix with (i,j)-entry

$$J(f)(x)_{i,j} = \begin{cases} 1 & \text{if } f_i(\overline{x}^j) \neq f_i(x), \\ 0 & \text{otherwise.} \end{cases}$$

A *signed directed graph* is a directed graph with a sign, $+1$ or -1, attached to each edge. Given $f : \{0,1\}^n \to \{0,1\}^n$ and $x \in \{0,1\}^n$, define

$$G(f)(x)$$

to be the signed directed graph with vertex set $\{1,\dots,n\}$ and with an edge from j to i when $J(f)(x)_{i,j} = 1$, with positive sign when

$$x_j = f_i(x),$$

and negative sign otherwise. A *signed edge* of a signed graph G is a triple (i,j,ε) such that G has an edge with sign ε from i to j. Such a triple will be denoted by $i \overset{\varepsilon}{\to} j$.

A *circuit* in a signed graph G is a non-empty sequence

$$k_1 \overset{\varepsilon_1}{\to} k_2 \overset{\varepsilon_2}{\to} \cdots \overset{\varepsilon_{p-1}}{\to} k_p \overset{\varepsilon_p}{\to} k_1$$

of signed edges of G. The *sign of a circuit* C is the product of the signs of its edges.

For instance, in the example of Figure 1 corresponding to $f(x) = (\overline{x_2}, \overline{x_1})$, it is easy to check that the Jacobian matrix associated to any state x is therefore given by:

$$J(x) = \begin{pmatrix} \overline{x_2} + \overline{x_2} & x_2 + \overline{x_2} \\ x_1 + \overline{x_1} & \overline{x_1} + \overline{x_1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

where the sum here is the sum of $\{0,1\}$ identified with the field $\mathbb{F}_2$. Therefore, the graph $G(x)$ at any state consists in a circuit between 1 and 2, hence a $\{1,2\}$-circuit. Since $x_1 \neq f_2(x)$ and $x_2 \neq f_1(x)$, the two edges are negative and the circuit is:

$$1 \overset{-1}{\longrightarrow} 2 \overset{-1}{\longrightarrow} 1, \quad \text{or simply} \quad 1 \dashv 2 \dashv 1,$$

with T-end notation for inhibitions, and is positive.

2.4 Functionality

The signed directed graph $G(f)(x)$ attached to each state x encompasses a subset of the regulatory interactions found in the complete regulatory network. These graphs are analogous to the local interaction graphs considered in [26] for instance. Consequently, in our discrete framework, a regulatory interaction and its sign may depend on the context, i. e., on the state of the system, in particular on the values of co-regulators acting on the same target. By taking unions of graphs on states x, we lose some details on the regulatory network and recover more global notions, closer to the objects usually manipulated by biologists: let $G(f) = \bigcup_{x \in \{0,1\}^n} G(f)(x)$ be the graph with a positive (resp. negative) edge from j to i when there exists $x \in \{0,1\}^n$ such that $G(f)(x)$ contains a positive (resp. negative) edge from j to i. Note that $G(f)$ may have both a positive and a negative edge between two given vertices.

This discussion motivates the following definition of the functionality context of a signed edge e: intuitively the set of states at which e is effective, or functional [18]. The functionality context of a circuit is then a notion of particular significance (as we shall see in Section 4). It is defined in the obvious way as follows.

DEFINITION 1 (Functionality context). *Let $f : \{0,1\}^n \to \{0,1\}^n$, $i,j \in \{1,\dots,n\}$, $\varepsilon \in \{+1,-1\}$, and let $e = (i,j,\varepsilon)$. The* functionality context $\Phi(f)(e)$ *of e is the set of $x \in \{0,1\}^n$ such that $G(f)(x)$ has an edge from i to j with sign ε. If C is a circuit, then $\Phi(f)(C) = \bigcap \Phi(f)(e)$ where e runs over signed edges of C. A circuit C is said to be* functional *when $\Phi(f)(C) \neq \varnothing$.*

Clearly, $x \in \Phi(f)(C)$ if and only if C is a circuit of $G(f)(x)$.

2.5 Globally minimal circuits

We shall be interested in a specific kind of circuits in regulatory graphs, namely circuits C occurring in some $G(f)(x)$, with the additional property that the global graph $G(f)$ has no other edge between vertices of C than the edges of C itself.

DEFINITION 2 (Minimal circuit). *Let Γ be a directed graph. The set of circuits of Γ is (partially) ordered as follows: if C_1, C_2 are circuits with vertex sets X_1, X_2 respectively, then $C_1 < C_2$ if and only if $X_1 \subsetneq X_2$. A circuit C is then said to be* minimal *when it is minimal for this order.*

DEFINITION 3 (Globally minimal circuit). *Let $f : \{0,1\}^n \to \{0,1\}^n$ and $x \in \{0,1\}^n$ such that $G(f)(x)$ contains a circuit C. We shall say that C is* globally minimal *if it is minimal as a circuit in $G(f)$.*

3 ISOLATED CIRCUITS

We reformulate the following result proved in [16]. According to the definition of the asynchronous dynamics, see (1), this result determines the dynamics of an isolated circuit, i. e., a regulatory graph constantly equal to a circuit.

THEOREM 1. *If $f : \{0,1\}^n \to \{0,1\}^n$ is such that for any $x \in \{0,1\}^n$, $G(f)(x)$ equals the circuit*

$$1 \overset{\varepsilon_1}{\to} 2 \overset{\varepsilon_2}{\to} \cdots \overset{\varepsilon_{n-1}}{\to} n \overset{\varepsilon_n}{\to} 1,$$

then for any $x \in \{0,1\}^n$, $f_i(x) \neq x_i$ if and only if

$$\begin{cases} x_{i-1} \neq x_i & \text{when } \varepsilon_{i-1} = +1, \\ x_{i-1} = x_i & \text{when } \varepsilon_{i-1} = -1, \end{cases}$$

if and only if $(-1)^{x_{i-1}+x_i} \neq \varepsilon_{i-1}$*, where indices are considered modulo* n *(i. e.,* $n+1=1$*) and the sum in the last inequality is the sum of the field* $\mathbb{F}_2$*.*

4 GLOBALLY MINIMAL CIRCUITS

Let us start with some notations. If $\kappa = x[\![I]\!]$ is a face of $\{0,1\}^n$, let $\pi_\kappa : \{0,1\}^n \to \kappa$ be the projection onto the affine subspace κ (identified with $\{0,1\}^I$), i. e., $\pi_\kappa(y)_i = y_i$ for any $i \in I$, and let $\sigma_\kappa : \kappa \to \{0,1\}^n$ be inclusion map of κ into $\{0,1\}^n$, i. e.,

$$\sigma_\kappa(y)_i = \begin{cases} y_i & \text{if } i \in I, \\ x_i & \text{otherwise.} \end{cases}$$

It is immediate that the definition of σ_κ does not depend on the choice of x such that $\kappa = x[\![I]\!]$, and that $\pi_\kappa \circ \sigma_\kappa$ is the identity. The folllowing Lemma, an equivalent simpler reformulation of Lemma 1 in [17], is a commutation property between the Jacobian and projection (or restriction).

LEMMA 1. *If* $f : \{0,1\}^n \to \{0,1\}^n$*,* $\kappa = x[\![I]\!]$ *is a face of* $\{0,1\}^n$ *and* $y \in \kappa$*, then:*

$$G(\pi_\kappa \circ f \circ \sigma_\kappa)(y) = G(f)(\sigma_\kappa(y))\restriction_I .$$

Proof — Let $i, j \in I$ and $y \in \kappa$. Since $i, j \in I$,

$$(\pi_\kappa \circ f \circ \sigma_\kappa)_j(\overline{y}^i) = f_j(\sigma_\kappa(\overline{y}^i)) = f_j\left(\overline{\sigma_\kappa(y)}^i\right).$$

Similarly, $(\pi_\kappa \circ f \circ \sigma_\kappa)_j(y) = f_j(\sigma_\kappa(y))$. Therefore,

$$(\pi_\kappa \circ f \circ \sigma_\kappa)_j(\overline{y}^i) \neq (\pi_\kappa \circ f \circ \sigma_\kappa)_j(y)$$

if and only if

$$f_j\left(\overline{\sigma_\kappa(y)}^i\right) \neq f_j(\sigma_\kappa(y)).$$

Moreover, since $i \in I$, $y_i = (\sigma_\kappa(y))_i$, and:

$$y_i + (\pi_\kappa \circ f \circ \sigma_\kappa)_j(y) = (\sigma_\kappa(y))_i + f_j(\sigma_\kappa(y)).$$

Consequently, signed edges in $G(\pi_\kappa \circ f \circ \sigma_\kappa)(y)$ and $G(f)(\sigma_\kappa(y))\restriction_I$ are the same. □

Then we show that the presence of a globally minimal circuit C has some important consequences on the dynamics restricted to the coordinates involved in C. Essentially, it enables to consider C as an isolated circuit.

THEOREM 2. *Let* $f : \{0,1\}^n \to \{0,1\}^n$*,* $x \in \{0,1\}^n$*, and suppose that* $G(f)(x)$ *contains a circuit*

$$C = k_1 \xrightarrow{\varepsilon_1} k_2 \xrightarrow{\varepsilon_2} \cdots \xrightarrow{\varepsilon_{p-1}} k_p \xrightarrow{\varepsilon_p} k_1$$

which is globally minimal. Let $\kappa = x[\![\{k_1, \ldots, k_p\}]\!]$*. Then* $\Phi(f)(C) \supseteq \kappa$ *and the dynamics of* $\pi_\kappa \circ f \circ \sigma_\kappa : \kappa \to \kappa$ *is given by Theorem 1.*

Proof — Let us first prove that $\Phi(f)(C) \supseteq \kappa$. To this end, let us consider $y \in \Phi(f)(C)$ and $i \in \{1, \ldots, p\}$ and let us show that $\overline{y}^{k_i} \in \Phi(f)(C)$. Since $y \in \Phi(f)(C)$, $G(f)(y)$ has a signed edge $(k_j, k_{j+1}, \varepsilon_j)$ for each $j \in \{1, \ldots, p\}$, i. e.:

$$f_{k_{j+1}}(y) \neq f_{k_{j+1}}(\overline{y}^{k_j})$$

and

$$\varepsilon_j = (-1)^{y_{k_j} + f_{k_{j+1}}(y)},$$

where indices are considered modulo p. Now, if $j = i$, it is straightforward that the signed edge $(k_j, k_{j+1}, \varepsilon_j)$ is in $G(f)(\overline{y}^{k_i})$ too; for the sign, simply observe that:

$$(\overline{y}^{k_j})_{k_j} + f_{k_{j+1}}(\overline{y}^{k_j}) = \overline{y_{k_j}} + \overline{f_{k_{j+1}}(y)} = y_{k_j} + f_{k_{j+1}}(y).$$

On the other hand, if $j \neq i$, since the circuit C is globally minimal, $G(f)$ has no signed edge from k_i to k_{j+1}, and in particular:

$$f_{k_{j+1}}(\overline{y}^{k_i}) = f_{k_{j+1}}(y) \tag{2}$$

and

$$f_{k_{j+1}}(\overline{y}^{k_i,k_j}) = f_{k_{j+1}}(\overline{y}^{k_j}),$$

therefore:

$$f_{k_{j+1}}(\overline{y}^{k_i}) \neq f_{k_{j+1}}(\overline{y}^{k_i,k_j})$$

and $G(f)(\overline{y}^{k_i})$ has an edge from k_j to k_{j+1}. Moreover, by (2) and $i \neq j$, we have:

$$(\overline{y}^{k_i})_{k_j} + f_{k_{j+1}}(\overline{y}^{k_i}) = y_{k_j} + f_{k_{j+1}}(y),$$

and the sign of this edge is ε_j. This holds for any $j \in \{1, \ldots, p\}$, and as a consequence, $\overline{y}^{k_i} \in \Phi(f)(C)$ when $y \in \Phi(f)(C)$ and $i \in \{1, \ldots, p\}$. Since $y \in \kappa \cap \Phi(f)(C)$, it follows that $\Phi(f)(C) \supseteq \kappa$.

Let us now prove that the dynamics of $\pi_\kappa \circ f \circ \sigma_\kappa$ satisfies this hypothesis of Theorem 1, i. e., that for any $y \in \kappa$, $G(\pi_\kappa \circ f \circ \sigma_\kappa)(y)$ equals the circuit C. By Lemma 1, it suffices to observe that $G(\pi_\kappa \circ f \circ \sigma_\kappa)(y)$ is the restriction of $G(f)(\sigma_\kappa(y))$ to vertices in I, and by the previous discussion, this coincides with C, q.e.d. □

We are now in position to combine Theorem 1 and Theorem 2 and delineate the dynamical properties implied by a globally minimal circuit.

THEOREM 3. *Under the hypotheses of Theorem 2, if* C *is positive, then* f *has two* $\{k_1, \ldots, k_p\}$*-fixed points; and if* C *is negative, then* f *has a* $\{k_1, \ldots, k_p\}$*-attractive cycle.*

Proof — If C is positive, by Theorem 1 and Theorem 2, $\pi_\kappa \circ f \circ \sigma_\kappa$ has two fixed points $P(0)$ and $P(1)$ defined by:

$$P(0)_{k_1} = 0,$$
$$P(1)_{k_1} = 1,$$
$$P(\alpha)_{k_i} \neq P(\alpha)_{k_{i+1}} \Leftrightarrow \varepsilon_i = -1, \quad \alpha = 0, 1, i = 1, \ldots, p-1.$$

Of course, $P(0)$ and $P(1)$ are fixed points of f because, by the positivity of C, $P(\alpha)_{k_1} \neq P(\alpha)_{k_p}$ if and only if $\varepsilon_p = -1$. Therefore, for each $\alpha = 0, 1$, $\sigma_\kappa(P(\alpha))$ and $f(\sigma_\kappa(P(\alpha)))$ have the same projection under π_κ. Hence, $\sigma_\kappa(P(0))$ and $\sigma_\kappa(P(1))$ are $\{k_1, \ldots, k_p\}$-fixed points.

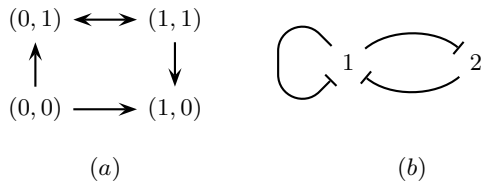

Fig. 3. (a) A perturbation of the dynamics of Figure 1. (b) The regulatory graph.

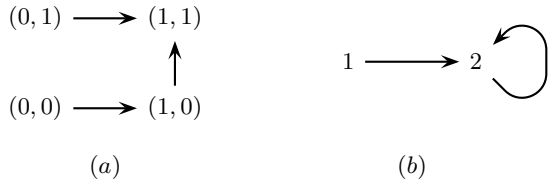

Fig. 4. (a) An example of dynamics with a globally minimal circuit (loop on 2), two 2-fixed points, but a single global fixed point. (b) The regulatory graph.

If C is negative, by Theorem 1 and Theorem 2, it is easy to check that $\pi_\kappa \circ f \circ \sigma_\kappa$ has an attractive cycle

$$P(0), \overline{P(0)}^{k_1}, \overline{P(0)}^{k_1,k_2}, \ldots, \overline{P(0)}^{k_1,\ldots,k_p} = P(1),$$
$$\overline{P(1)}^{k_1}, \overline{P(1)}^{k_1,k_2}, \ldots, \overline{P(1)}^{k_1,\ldots,k_p} = P(0).$$

Hence the image

$$\sigma_\kappa(P(0)), \sigma_\kappa\left(\overline{P(0)}^{k_1}\right), \ldots$$

of this cycle under σ_κ is a $\{k_1, \ldots, k_p\}$-attractive cycle of f. □

The global minimality hypothesis in Theorems 2 and 3 cannot be simply avoided. For instance, the dynamics corresponding to the map $f : \{0,1\}^2 \to \{0,1\}^2$ defined by $f(x) = (\overline{x_2}, \overline{x_1})$ gives rise to a globally minimal positive circuit and indeed has two fixed points $(0,1)$ and $(1,0)$ (Figure 1), whereas the perturbated dynamics corresponding to $g(x) = (\overline{x_1 x_2}, \overline{x_1})$ has a single fixed point $(1,0)$: the $\{1,2\}$-circuit is no more globally minimal, it is perturbated by the negative loop on 1 (Figure 3).

It is not true either that the localised dynamics predicted by the above results leads necessarily to the corresponding global behaviour. In particular, the presence of a globally minimal positive circuit does not imply the existence of disjoint stable subspaces in general. This can be seen by considering the map $h(x) = (1, x_1 \vee x_2)$. The positive circuit consisting in a loop on 2 is globally minimal and its functionality context is given by $x_1 = 0$. The dynamics, which is given in Figure 4, has two 1-fixed points $(0,0)$ and $(0,1)$, but the only global fixed point of h is $(1,1)$: the positive loop on 2 acts as a "partial separator" between the subspaces $x_2 = 0$ and $x_2 = 1$. A natural question is therefore to understand more precisely under which conditions these modules combine to produce global separators and global differentiation.

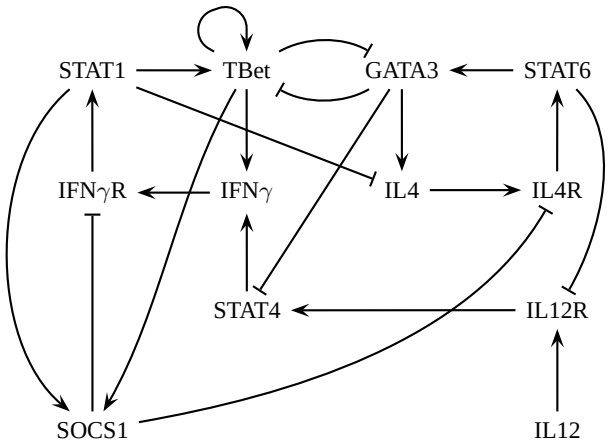

Fig. 5. Regulatory graph of the network controlling Th lymphocyte differentiation. The nodes represent transcription regulatory factors (Tbet, GATA3), signaling transduction factors (STAT1, STAT4, STAT6, SOCS1), lymphokines (IFNγ, IL4, IL12) and receptors (IFNγR, IL4R, IL12R). Remark that IL12 acts as an input of the system.

5 APPLICATION

We present here a biological illustration and then apply the results proved in the previous Section.

We consider the network involved in the control of the Th-lymphocyte differentiation. The vertebrate immune system contains various cell populations. Among B and T lymphocytes, CD4+ T helper lymphocytes can further differentiate into T-helper 1 (Th1) or Th2 cells, which respectively enable cell mediated immunity and humoral responses. Th1 and Th2 cells can be distinguished according to their pattern of cytokine secretion. Immune responses biased towards the Th1 phenotype result in autoimmune diseases, while enhanced Th2 responses originate allergic reactions [1, 13]. Various mathematical models have been proposed for the differentiation, activation and proliferation of Th-lymphocytes. Many of them were focusing on interactions between immunological cell populations at a macroscopic level [4, 33, 34]. Other model analyses aim at understanding the mechanism of the generation of antibody and T-cell receptorsdiversity, as well as the dynamical properties of the large networks defined by the interactions between cytokines [10] or between immunoglobulins (see, e.g., [32]). We consider here a very simplified Boolean modelling of this Th-lymphocyte differentiation already presented in [18], which involves 12 regulatory components (Figure 5). Other regulatory graphs using the same discrete modelling (Boolean or multivalued) have been proposed [12, 14].

It has been shown [11] that the system can reach the three stable states given in Table 1. The first stable state s_1 corresponds to the virgin Th cells (Th0), whereas the second and third ones s_2, s_3 correspond respectively to Th2 and Th1 differentiated lymphocytes.

5.1 Functional circuits

The regulatory graph represented in Figure 5 contains 18 circuits. Only 4 of them are functional, in the sense of Definition 1. Among

Genes		IFNγ	IL4	IL12	IFNγR	IL4R	IL12R	STAT1	STAT6	STAT4	SOCS1	Tbet	GATA3
Stable states	s_1	0	0	0	0	0	0	0	0	0	0	0	0
	s_2	0	1	0	0	1	0	0	1	0	0	0	1
	s_3	1	0	0	0	0	0	0	0	0	1	1	0

Table 1. The three stable states s_1, s_2, s_3, which represent respectively the naive, Th2 and Th1 cells.

these functional circuits, three are positive:

$$C1 = \left(\text{IL4R} \xrightarrow{+} \text{STAT6} \xrightarrow{+} \text{GATA3} \xrightarrow{+} \text{IL4} \xrightarrow{+} \text{IL4R}\right),$$

$$C2 = \left(\text{Tbet} \xrightarrow{+} \text{Tbet}\right),$$

$$C3 = \left(\text{GATA3} \xrightarrow{-} \text{Tbet} \xrightarrow{-} \text{GATA3}\right),$$

and one is negative:

$$C4 = \left(\text{IFN}\gamma\text{R} \xrightarrow{+} \text{STAT1} \xrightarrow{+} \text{SOCS1} \xrightarrow{-} \text{IFN}\gamma\text{R}\right).$$

Let $f : \{0,1\}^{12} \to \{0,1\}^{12}$ be the map corresponding to the asynchronous dynamics (not shown here for sake of space). The graph $G(f)$ represented Figure 5 is the union of all the local graphs $G(f)(x)$ for $x \in \{0,1\}^{12}$. Only $C1, C2$ and $C4$ are globally minimal, $C3$ is not because of the loop $C2$. Let us compute the functionality contexts of these circuits.

- Circuit $C1$ is functional when Tbet, STAT1 and SOCS1 are not expressed, therefore $\Phi(f)(C1) = \{x \mid x_{\text{Tbet}} = x_{\text{STAT1}} = x_{\text{SOCS1}} = 0\}$.
- Circuit $C2$ (self-regulation of Tbet) is functional when STAT1 and GATA3 are not expressed, i. e., $\Phi(f)(C2) = \{x \mid x_{\text{GATA3}} = x_{\text{STAT1}} = 0\}$.
- The non globally minimal circuit $C3$ is functional when STAT6 and STAT1 are expressed, i. e., $\Phi(f)(C3) = \{x \mid x_{\text{STAT6}} = x_{\text{STAT1}} = 1\}$.
- Finally, the negative circuit $C4$ is functional when Tbet is not expressed and IFNγ expressed, i. e., $\Phi(f)(C4) = \{x \mid x_{\text{IFN}\gamma} = 1,\ x_{\text{Tbet}} = 0\}$.

Note that the functionality contexts of $C1$ and $C2$ are compatible and overlap: they both require the absence of STAT1. On the other hand, when STAT1 is expressed, circuit $C3$ is functional.

5.2 Analysis and comments

Let us consider the circuit $C1$. By Theorem 2, we know the structure of the states space of any face $x[\![\{\text{IL4R}, \text{STAT6}, \text{GATA3}, \text{IL4}\}]\!]$ with $x \in \Phi(f)(C1)$. Moreover, by Theorem 3, as $C1$ is positive, there are two $\{\text{IL4R}, \text{STAT6}, \text{GATA3}, \text{IL4}\}$-fixed points. Here, s_1 and s_2 belong to $\Phi(f)(C1)$, and they differ only in the four coordinates which correspond to the four genes of $C1$: $s_1[\![\{\text{IL4R}, \text{STAT6}, \text{GATA3}, \text{IL4}\}]\!] = s_2[\![\{\text{IL4R}, \text{STAT6}, \text{GATA3}, \text{IL4}\}]\!]$. These two local fixed points are also stable for the whole dynamics.

We can do the same type of analysis for the circuit $C2$. Theorem 2 gives the structure of the states space of any face $x[\![\{\text{Tbet}\}]\!]$ with $x \in \Phi(f)(C2)$. Moreover, as $C2$ is positive, there are two $\{\text{Tbet}\}$-fixed points (Theorem 3). Here again, two of the three global fixed points belong to the context of functionality of $C2$: $s_1, s_2 \in \Phi(f)(C2)$.

When Tbet is not expressed (for example by the indirect effect of a perturbation of IL4, as proposed in [12]), GATA3 can be activated, and the circuit $C1$ is functional. Hence, the system reaches the differentiated state s_2 (which represents Th2 cells). But if the expression of Tbet increases, for example because the lymphokines IFNγ is transiently expressed, then $C1$ is no more functional, but $C2$ is, and this self-regulation maintains Tbet expressed. Then, the system reaches the differentiated state Th1 (s_3).

Concerning the negative circuit $C4$, by Theorems 2 and 3, we know that any face $x[\![\{\text{IFN}\gamma\text{R}, \text{STAT1}, \text{SOCS1}\}]\!]$ with $x \in \Phi(f)(C4)$ has a $\{\text{IFN}\gamma\text{R}, \text{STAT1}, \text{SOCS1}\}$-attractive cycle. In fact, the dynamics restricted to $\Phi(f)(C4)$, i. e., the restriction of f to $\bigcup_{x \in \Phi(f)(C4)} x[\![\{\text{IFN}\gamma\text{R}, \text{STAT1}, \text{SOCS1}\}]\!]$, contains an attractive cycle, where all the genes not in $C4$ and $\Phi(f)(C4)$ are not expressed. The negative circuit $C4$ is functional when the lymphokine IFNγ is expressed and Tbet is not expressed. This functionality context is therefore fragile: as Tbet is an activator of IFNγ, the absence of Tbet implies that the expression level of IFNγ tends to 0, hence $C4$ should stay functional for a short time.

As it is proved in Section 4, when a circuit is considered in a state which belongs to its functionality context, then, letting only the variables of the circuit free, the structure of the dynamics is the same as the one of an isolated circuit. Hence, we have a precise local knowledge of the dynamics.

In this application, the functionality contexts of the 3 positive circuits cover all the phase space. Each positive circuit creates in its functionality context 2 basins of attraction, and finally, the whole space is divided into 3 basins corresponding to the 3 stable states (their maximal number is 2^3 in general [3]). Therefore, one of the challenges is now to be able to describe more precisely the position of the basins of attraction, where they separate and how they possibly connect each other.

6 CONCLUSION

Even when the dynamics, i. e., the function f is known, the study of the phase space is not easy, and often not computationally feasible. The idea of getting as more information as possible on the dynamics from the structure of the regulatory graph —which is much smaller— is really attractive. The property of functionality of a circuit is well suited for this purpose. Indeed, the important role of the circuits on the dynamics of the systems is well-known, but the number of circuits in a typical regulatory graph is generally quite large. Fortunately, the circuits which have a real incidence on the dynamics are the functional ones, at least the globally minimal functional ones according to Section 4, and their number is much more accessible. For instance, in the illustration considered in Section 5, the model contains 18 circuits, but only 4 of them are functional. Actually, the proportion of functional circuits can be often much smaller in practice.

An interesting current line of research is therefore to decompose regulatory graphs into modules. The notion of modularity is not

trivial, and the definition of a module is not straightforward, and, at least, not unique. This article leads to naturally define modules around the notion of globally minimal functional circuits.

While the example studied in Section 5 provides a good illustration of the potentiality of the method, the present work is clearly in progress, and many improvements can certainly be done. The generalisation of our results to multivalued dynamics, as in [19], requires a careful definition of regulatory graphs and functionality of circuits. The possibility of a sufficient condition on the Jacobian matrix of differential or piecewise-linear systems [27, 8, 25] is worth exploring too.

On the other hand, relaxing the minimality constraint on circuits in Theorems 2 and 3 seems to require further work. The presence of a non minimal functional circuit is indeed not so rare. For instance, the self-regulations of "clue-genes", involved in functional circuits, should create this situation. Therefore, this constraint prevents us from analysing some important circuits.

REFERENCES

[1]D. Agnello, C.S.R. Lankford, J. Bream, A. Morinobu, M. Gadina, J. O'Shea, and D.M. Frucht. Cytokines and transcription factors that regulate t helper cell differentiation: new players and new insights. *J. Clin. Immunol.*, 23, 2003.

[2]R. Albert and H. Othmer. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *drosophila melanogaster*. *J.Theoret. Biol.*, 223:1–18, 2002.

[3]J. Aracena, J. Demongeot, and E. Goles. On limit cycles of monotone functions with symmetric connection graph. *Theoretical Computer Science*, 322(2):237–244, 2004.

[4]C. Bergmann, J.L. Van Hemmen, and L.A. Segel. Th1 or th2: how an appropriate t helper response can be made. *Bulletin of Mathematical Biology*, 63:405–430, 2001.

[5]H. de Jong. Modeling and simulation of genetic regulatory systems: A literature review. *Journal of Computational Biology*, 9(1):67–103, 2002.

[6]A. Ghysen and R. Thomas. The formation of sense organs in Drosophila: a logical approach. *BioEssays*, 25:802–807, 2003.

[7]L. Glass and S. A. Kauffman. The logical analysis of continuous non-linear biochemical control networks. *Journal of Theoretical Biology*, 39:103–129, 1973.

[8]J.-L. Gouzé. Positive and negative circuits in dynamical systems. *Journal of Biological Systems*, 6:11–15, 1998.

[9]S. A. Kauffman. *The origins of order: Self-organization and selection in evolution*. Oxford University Press, 1993.

[10]G. Krueger, G. Marshall, U. Junker, H. Schroeder, L. Buja, and G. Wang. Growth factors, cytokines, chemokines and neuropeptides in the modeling of t-cells. *In vivo*, 16, 2002.

[11]L. Mendoza. A network model for the control of the differentiation process in Th cells. *Biosystems*, 84(2), 2006.

[12]L. Mendoza and I. Xenarios. A method for the generation of standardized qualitative dynamical systems of regulatory networks. *Theoret. Biol. Med. Mod.*, 3(13):1–18, 2006.

[13]K.M. Murphy and S.L. Reiner. The lineage decisions on helper t cells. *Nature Reviews Immunology*, 2:933–944, 2002.

[14]A. Naldi, D. Thieffry, and C. Chaouiya. Decision Diagrams for the Representation and Analysis of Logical Models of Genetic Networks. *CMSB 2007, LNBI* 4695:233–247, 2007.

[15]E. Plahte, T. Mestl, and S. W. Omholt. Feedback loops, stability and multistationarity in dynamical systems. *Journal Biological Systems*, 3:409–413, 1995.

[16]É. Remy, B. Mossé, C. Chaouiya, and D. Thieffry. A description of dynamical graphs associated to elementary regulatory circuits. *Bioinformatics*, 19(2):172–178, 2003.

[17]É. Remy and P. Ruet. On differentiation and homeostatic behaviours of Boolean dynamical systems. In *Transactions on Computational Systems Biology*, volume 4780 of *Lecture Notes in Computer Science*, pages 92–101. Springer, 2007.

[18]É. Remy, P. Ruet, L. Mendoza, D. Thieffry and C. Chaouiya. From logical regulatory graphs to standard Petri nets: dynamical roles and functionality of feedback circuits. In *Trans. Comput. Systems Biol.*, vol. 4230 of *LNCS*, pp. 56–72, 2006.

[19]É. Remy, P. Ruet, and D. Thieffry. Positive or negative regulatory circuit inference from multilevel dynamics. In *Positive Systems: Theory and Applications*, volume 341 of *LNCIS*, pages 263–270. Springer, 2006.

[20]É. Remy, P. Ruet, and D. Thieffry. Graphic requirements for multistability and attractive cycles in a Boolean dynamical framework. *Advances in Applied Mathematics*, To appear, 2008.

[21]F. Robert. *Discrete iterations: a metric study*, volume 6 of *Series in Computational Mathematics*. Springer, 1986.

[22]L. Sánchez and D. Thieffry. Segmenting the fly embryo: a logical analysis of the pair-rule cross-regulatory module. *Journal of Theoretical Biology*, 224:517–537, 2003.

[23]T. Schlitt and A. Brazma. Current approaches to gene regulatory network modelling. *BMC Bioinformatics*, 8(Suppl. 6):S9, 2007.

[24]I. Shmulevich, E. R. Dougherty, and W. Zhang. From Boolean to probabilistic Boolean networks as models of genetic regulatory networks. *Proc. IEEE*, 90(11):1778–1792, 2002.

[25]E. H. Snoussi. Necessary conditions for multistationarity and stable periodicity. *Journal of Biological Systems*, 6:3–9, 1998.

[26]C. Soulé. Graphic requirements for multistationarity. *ComPlexUs*, 1:123–133, 2003.

[27]C. Soulé. Mathematical approaches to gene regulation and differentiation. *CRAS : Biologies*, 329(1):13–20, 2006.

[28]D. Thieffry. Dynamical roles of biological regulatory circuits. *Briefings in Bioinformatics*, 8:220–225, 2007.

[29]R. Thomas. Boolean formalization of genetic control circuits. *Journal of Theoretical Biology*, 42:563–585, 1973.

[30]R. Thomas. On the relation between the logical structure of systems and their ability to generate multiple steady states and sustained oscillations. In *Ser. Synergetics* 9:180–193, 1981.

[31]R. Thomas, D. Thieffry, and M. Kaufman. Dynamical behaviour of biological regulatory networks I. Biological role of feedback loops and practical use of the concept of the loop-characteristic state. *Bulletin of Mathematical Biology*, 57:247–276, 1995.

[32]G. Weisbuch, R.J. De Boer, and A.S. Perelson. Localized memories in idiotypic networks. *Journal of Theoretical Biology*, 146:483–499, 1990.

[33]A. Yates, C. Bergman, J.L. Van Hemmen, J. Stark, and R. Callard. Cytokine-mediated regulation of helper t cell populations. *J. Theoret. Biology*, 206:539–560, 2000.

[34]A. Yates, J. Stark, N. Klein, R. Antia and R. Callard. Understanding the Slow Depletion of Memory CD4+ T Cells in HIV Infection. *PLoS Medecine*, 4(5): e177, 2007.